CONTINUOUS TURBIDITY MONITORING AND REGRESSION ANALYSIS TO ESTIMATE TOTAL SUSPENDED SOLIDS AND FECAL COLIFORM BACTERIA LOADS IN REAL TIME

By Victoria G. Christensen, Andrew C. Ziegler, and Xiaodong Jian
Hydrologists, U.S. Geological Survey, 4821 Quail Crest Place, Lawrence, KS 66049

Abstract: *To obtain timely and continuous water-quality information, the U.S. Geological Survey, in cooperation with State and other Federal agencies, has been using an innovative real-time monitoring approach for several Kansas streams. Continuously recorded data and data from periodic collection of water-quality samples are being used to develop surrogate relations between turbidity and constituents of concern. Regression equations were developed to estimate total suspended solids and fecal coliform bacteria from continuous turbidity measurements collected from 1995 through 1998 on the Little Arkansas River in south-central Kansas. The equations were applied to data collected in 1999 to estimate constituent loads. Estimated total suspended solids loads were 460 and 613 million pounds per year for the two sites in the study. Estimated fecal coliform bacteria loads were 30,000,000 and 32,000,000 billions of colonies per year. Despite some large differences between instantaneous measured and regression-estimated loads, continuous monitoring of turbidity in streams may increase the accuracy of load estimates that may be useful for calculating total maximum daily loads (TMDL's). The availability of this data in real time also may be useful for considering whole-body contact and recreation criteria, for adjusting water-treatment strategies, and for preventing adverse effects on fish or other aquatic life.*

INTRODUCTION

Historically, the U.S. Geological Survey's (USGS) stream-gaging network has provided timely water-quantity information to resource managers and others to make informed decisions about floods and water availability. It has not been possible, however, to provide water-quality information in the same timely manner. Timely water-quality information is useful for many reasons, including assessment of the effects of urbanization and agriculture on a water supply.

Nationally, the U.S. Environmental Protection Agency (USEPA) lists both suspended solids and bacteria as primary water-quality concerns. Suspended solids can cause problems for fish by clogging gills and for aquatic plants by reducing light penetration and thus limiting growth. In addition, suspended solids provide a medium for the accumulation and transport of other constituents such as phosphorus and bacteria. The presence of fecal coliform bacteria in surface water indicates fecal contamination and possibly the presence of other organisms that could cause disease.

In the past, to determine concentrations of total suspended solids and fecal coliform bacteria in a stream, it was necessary to manually collect samples and send them to a laboratory for analysis. This procedure took time, and when human health is a concern, immediate information is necessary. In addition, because manually collected samples did not provide continuous data, constituent loads during peak flows often were missed, making accurate load estimates difficult. Load estimates are important to the establishment and monitoring of total maximum daily loads (TMDL's) mandated by the Clean Water Act of 1972.

In response to the need for timely and continuous water-quality information, the USGS, in cooperation with State and other Federal agencies, has been using an innovative continuous, real-time monitoring approach for several Kansas streams. This paper describes results of a study to provide continuous estimates of real-time total suspended solids and fecal coliform bacteria loads for the Little Arkansas River, which is used as a source water for artificial recharge of the *Equus* Beds aquifer. The *Equus* Beds aquifer provides approximately 50 percent of the drinking water for the city of Wichita in south-central Kansas.

METHODS

To assess the quality of the Little Arkansas River, in-stream water-quality monitors were installed in 1998 at two U.S. Geological Survey stream-gaging stations near Halstead and Sedgwick, Kansas (fig. 1), to provide continuous, real-time measurement of turbidity and other physical properties of water. Periodic water samples were collected manually from 1995 through 1998 at the two gaging stations and analyzed for selected constituents and physical properties including turbidity, total suspended solids (TSS), and fecal coliform bacteria.

The manual samples were collected throughout the year and throughout 95 percent of the stream's typical flow duration to describe a wide range of seasonal and hydrologic conditions. Sampling techniques are described in

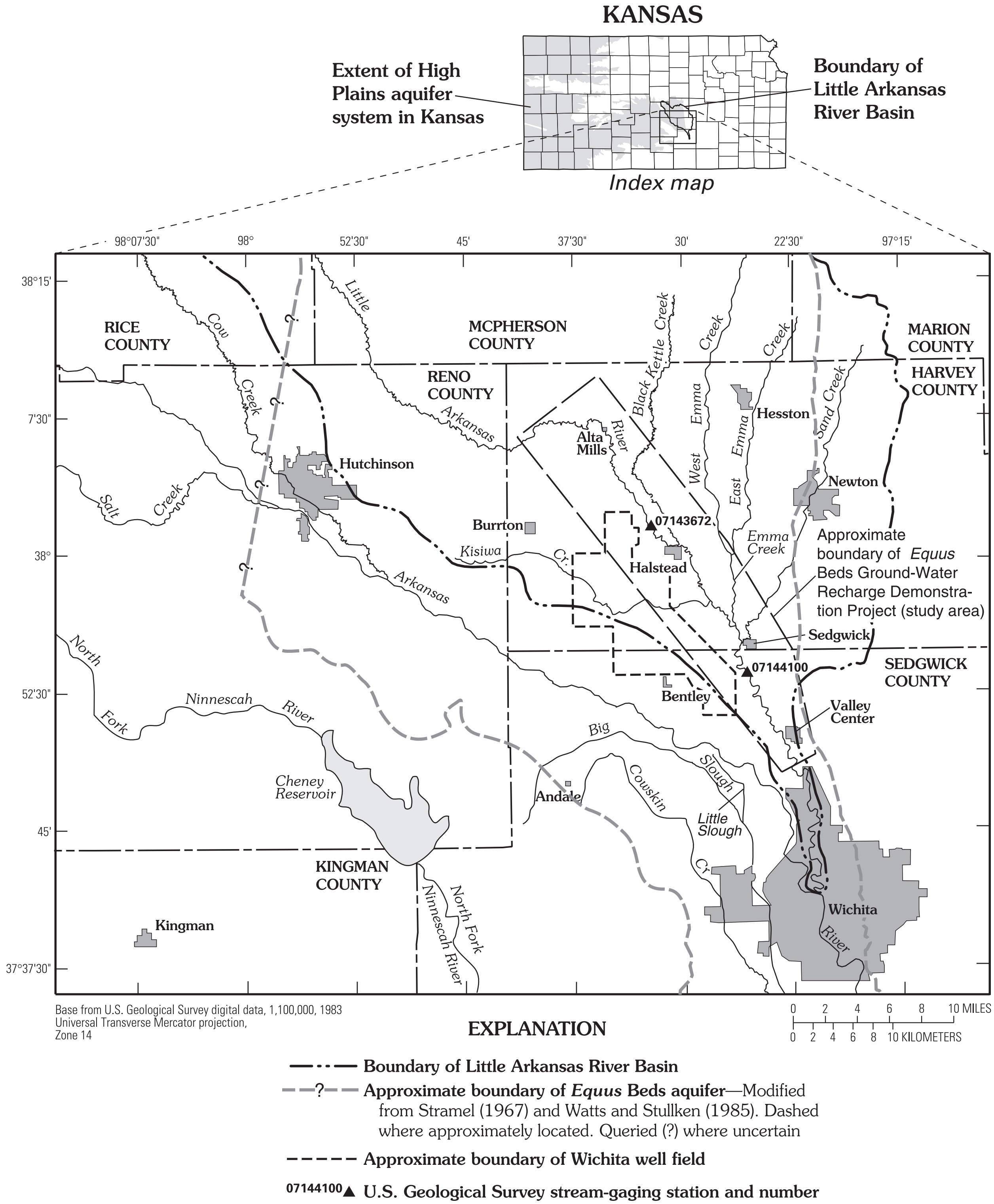

Figure 1. Location of study area in south-central Kansas.

Ziegler and Combs (1997). Linear regression equations were developed using the least-squares method.

To test the regression equations developed from the first 3 years of data collection (1995-98), the equations were applied to the fourth year of data collection (1999) to calculate estimated instantaneous constituent loads and errors associated with these loads. Additional information on methods related to this study can be found in Ziegler and Combs (1997), Ziegler and others (1999), and Christensen and others (2000). Current information on continuous, real-time water quality in Kansas may be accessed through the World Wide Web at http://ks.water.usgs.gov/Kansas/qw.

RESULTS

Turbidity Measurement: Turbidity can be an indicator of the amount of sediment and related constituents transported by a stream. Turbidity and streamflow are related (fig. 2) because streamflow can affect suspension of the sediment and related constituents. Continuous and periodic monitoring enable identification of seasonal trends and effects of extreme hydrologic conditions on turbidity, TSS, and fecal coliform bacteria and estimation of chemical loads transported in the Little Arkansas River. High flows have a significant effect on chemical loads, and concentration data from manual samples often are not available. Therefore, continuous monitoring of turbidity for the estimation of TSS and fecal coliform bacteria in streamflow may increase the accuracy of load estimates. Manual samples were collected using depth- and width-integrating techniques, and because a real-time water-quality monitor measures the turbidity at a single point in the stream, there was some difference between turbidity of water samples

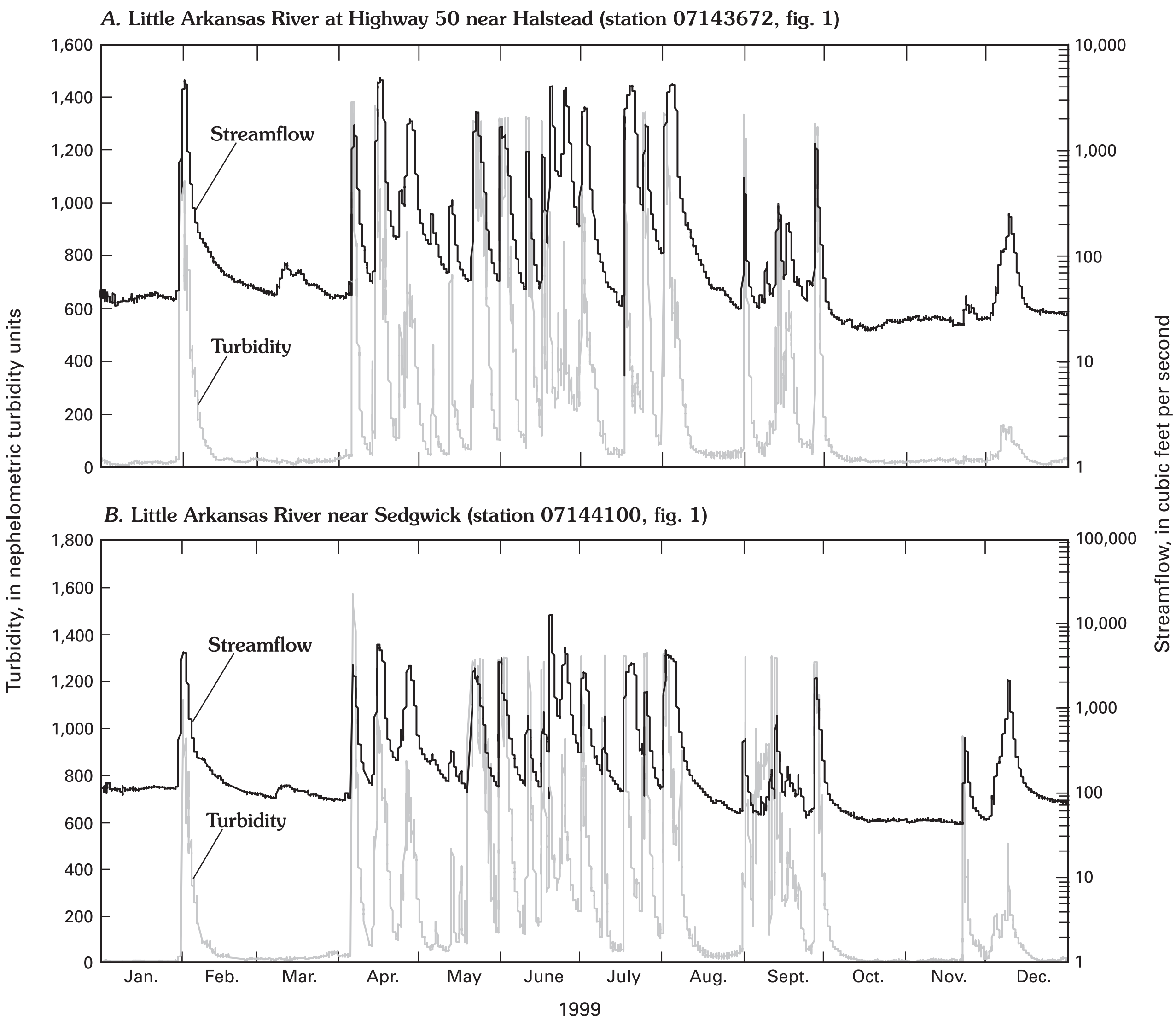

Figure 2. Comparison of turbidity and streamflow for Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

collected manually across the width and depth of the stream and turbidity from the real-time water-quality monitor. Turbidity measurements were not adjusted to account for this small difference.

Total Suspended Solids: Total suspended solids (TSS) include both suspended sediment and organic material collected with the water sample. The significance of organic material with respect to the determination of sediment concentration is minimal in most drainage basins (Guy, 1969). In this study, analysis for suspended sediment began in January 1999 and was not used to develop the regression equations. However, regression analysis of samples collected in 1999 indicate that the relation between *TSS* and total suspended sediment is highly correlated (V. Christensen, USGS, written commun., 2000).

At the Halstead gaging station there is a linear relation between TSS and turbidity after logarithmic transformation. Graphic plots of this relation indicate that there are two outliers (the minimum and maximum turbidity values of 0.3 and 1,780 nephelometric turbidity unit (NTU), respectively). To decrease the effects of these outliers, these two points were not used in the regression equation. The range in turbidity values used to develop the equation for the Halstead gaging station was 4.69 to 1,150 NTU. The final linear regression equation for Halstead is:

$$log_{10}(TSS) = 0.920 log_{10}(Turb) + 0.243, \tag{1}$$

where *TSS* is estimated total suspended solids, in milligrams per liter, and *Turb* is turbidity, in nephelometric turbidity units. The mean square error (MSE) for the equation is 0.193, and the coefficient of determination (R^2) is 0.911.

At the Sedgwick gaging station there were three outliers with turbidity values of 3.99, 5.01, and 1.44 NTU. The range in turbidity values used to develop the equation was 3.63 to 1,030 NTU. The final linear regression equation after logarithmic transformation for Sedgwick is:

$$log_{10}(TSS) = 0.878 log_{10}(Turb) + 0.300, \tag{2}$$

where *TSS* is estimated total suspended solids, in milligrams per liter, and *Turb* is turbidity, in nephelometric turbidity units. The MSE for the Sedgwick equation is 0.209, and the R^2 is 0.889.

The relation between TSS and turbidity is similar for the Halstead and Sedgwick gaging stations. In addition, it appears there is very little improvement to the relation with the addition of samples. For example, 19 samples collected throughout the range in hydrologic conditions may be sufficient to define the relation between TSS and turbidity (table 1). Changes in TSS load generally are affected by changes in streamflow (Topping and others, 2000), and thus, load shows a seasonal fluctuation (fig. 3).

The median relative percentage differences between measured (calculated from manual samples) and estimated TSS loads were 66.4 and 34.0 percent at the Halstead and Sedgwick gaging stations, respectively (Christensen and others, 2000). There was a large difference between estimated TSS loads at the upstream (Halstead) and downstream (Sedgwick) gaging stations (460 million and 613 million pounds in 1999, respectively) as expected. Turbidity and TSS may be altered significantly between the Halstead and Sedgwick gaging stations due to inflow of intervening drainages, Kisiwa, Emma, and Sand Creeks (fig. 1). Although the 1999 estimated TSS load was larger at the Sedgwick gaging station, the Halstead gaging station had the larger estimated TSS yield (1,050 pounds per acre compared to 826 pounds per acre). The relatively small load contribution of the intervening drainage area (fig. 4) may be because of storage of solids and sediment in the streambed between the Halstead and Sedgwick gaging stations or because the inflow of Kisiwa, Emma, and Sand Creeks may have a dilution effect on TSS at the Sedgwick gaging station.

Fecal Coliform Bacteria: Total coliform bacteria was identified by Ziegler and others (1999) as a constituent of concern in the Little Arkansas River during high flows. The equations in this paper are based on fecal coliform bacteria, one of the bacteria included in a total coliform analysis. Fecal coliform bacteria analyses were chosen for inclusion in this paper because current (2000) State of Kansas water-quality criteria [2,000 col/100 mL (colonies per 100 milliliters of water) for noncontact recreation, 200 col/100 mL for whole-body contact recreation, and less than 1 col/100 mL for drinking water] are based on fecal coliform bacteria densities (Kansas Department of Health and Environment, 1999). Because runoff from a watershed may transport fecal coliform bacteria to streams, there may be

Table 1. Sample-size effect on improving sum of squares of error (SSE) for the surrogate-based total suspended solids equation

[R^2, coefficient of determination; SSE, sum of the squares of error; %, percent; --, not determined]

Calendar year	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 1)				Little Arkansas River near Sedgwick (station 07144100, fig. 1)			
	Total number of samples	R^2	SSE	Change in SSE (%)	Total number of samples	R^2	SSE	Change in SSE (%)
1995	19	0.907	2.79	--	19	0.881	3.24	--
1996	41	.908	2.78	-0.36	35	.879	3.28	1.23
1997	58	.909	2.75	-1.08	51	.885	3.12	-4.88
1998	74	.911	2.69	-2.18	71	.889	3.01	-3.53

a relation between bacteria densities and streamflow or possibly to time of year because runoff characteristics may vary with season. However, only month (time) and turbidity were found to be significantly related to fecal coliform bacteria.

The range in turbidity values used in the development of the equation for fecal coliform bacteria at the Halstead gaging station was 0.3 to 1,780 NTU. The multiple regression equation with logarithmic transformation for Halstead is:

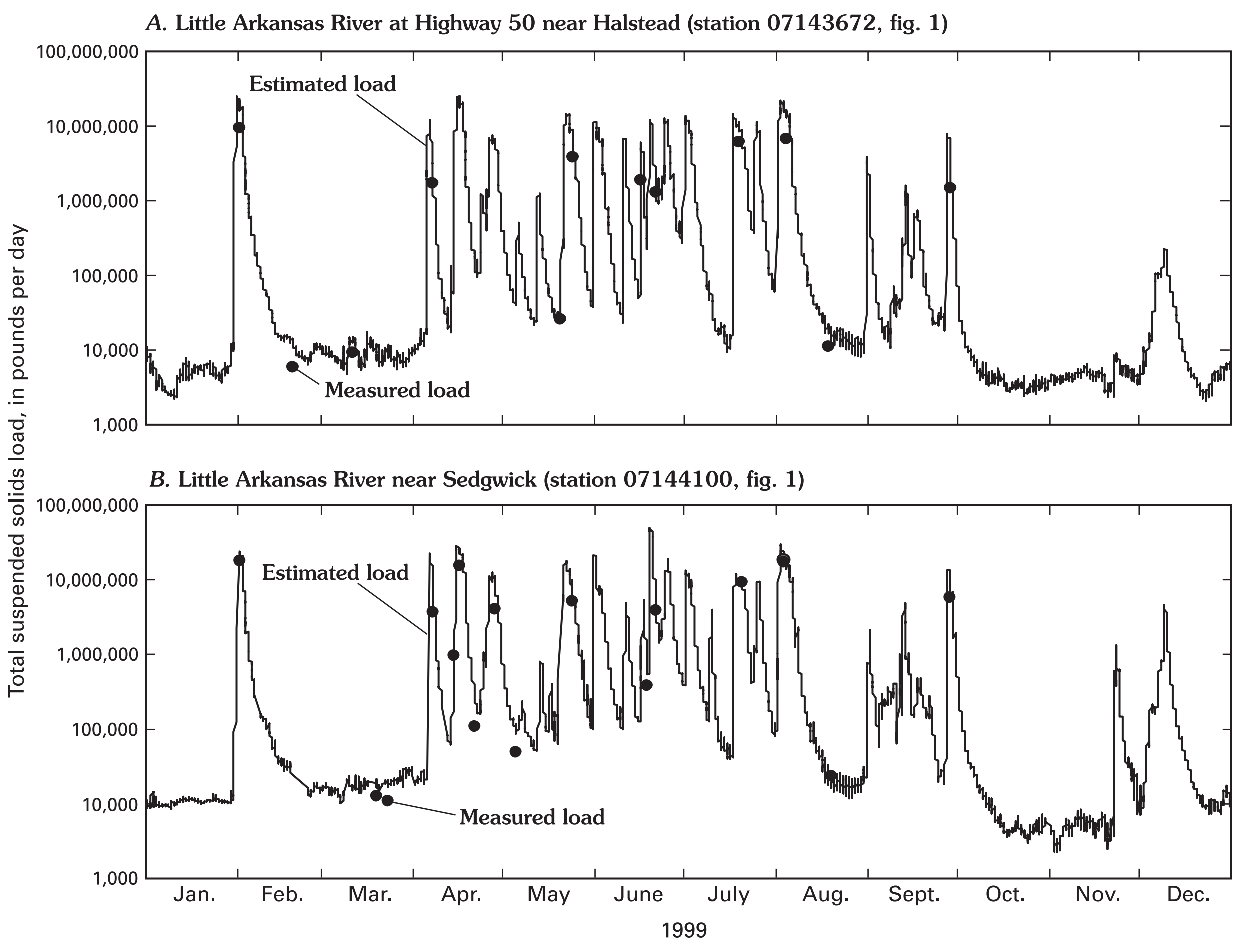

Figure 3. Comparison of measured and estimated total suspended solids load for Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

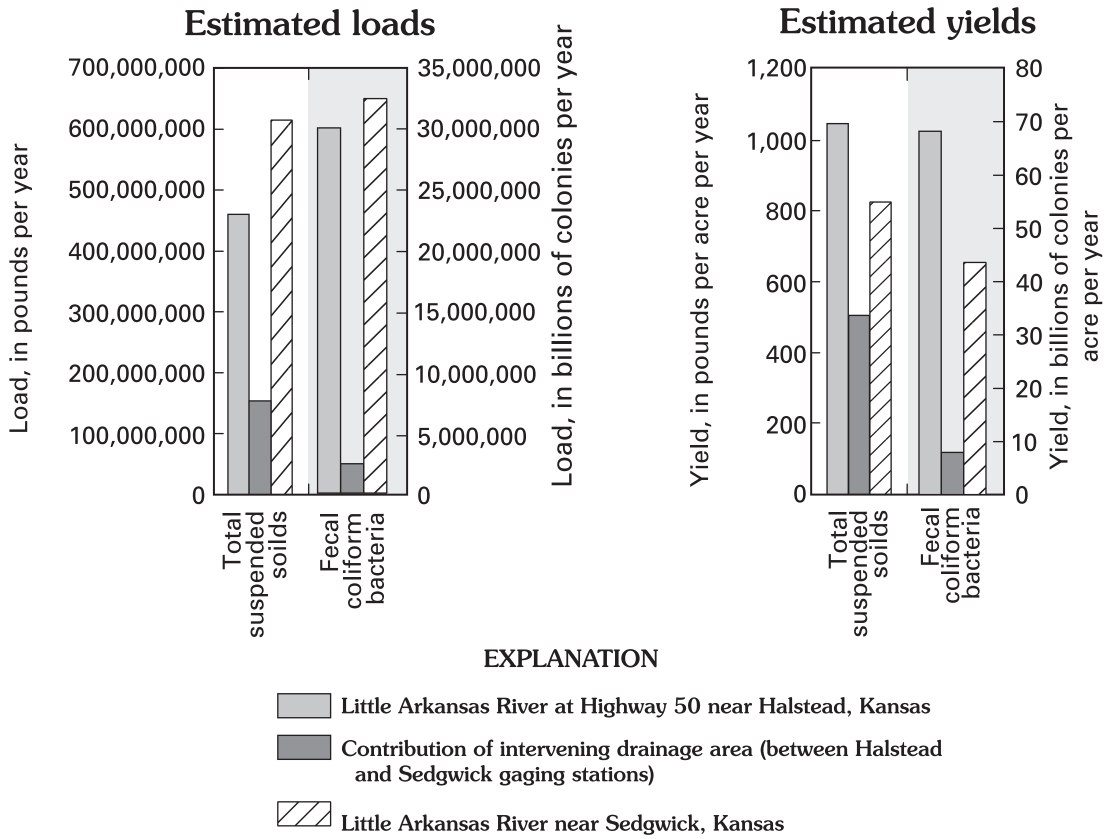

Figure 4. Comparison of estimated loads and yields for total suspended solids and fecal coliform bacteria for Little Arkansas River at Highway 50 near Halstead and near Sedgwick, Kansas, 1999.

$$\log_{10}(Bact) = 0.490\cos\left(2\pi\left(\frac{month + 2.06}{8.76}\right)\right) + 0.00106Turb + 0.417\log_{10}(Turb) + 1.65\,, \quad (3)$$

where *Bact* is fecal coliform bacteria density, in colonies per 100 milliliters of water; *month* is a number from 1 to 12; and *Turb* is turbidity, in nephelometric turbidity units. This equation has an MSE of 0.609 and R^2 of 0.620.

For the Little Arkansas River near Sedgwick, the range in turbidity values used in the development of the regression equation for fecal coliform bacteria was 1.44 to 1,030 NTU,   The final multiple regression equation for fecal coliform bacteria with logarithmic transformation for Sedgwick is:

$$\log_{10}(Bact) = -0.286\cos\left(2\pi\left(\frac{month - 29.4}{-6.8}\right)\right) - 0.000422(Turb) + 1.26\log_{10}(Turb) + 0.519\,, \quad (4)$$

where *Bact* is fecal coliform bacteria density, in colonies per 100 milliliters of water; *month* is a number from 1 to 12; and *Turb* is turbidity, in nephelometric turbidity units. This equation has an MSE of 0.784 and R^2 of 0.556.

The periodic cosine function is used in the bacteria equations for both stations with the period equal to 8.76 and 6.8 months, respectively. The physical basis for this period or cycle is related to the seasonal variability of fecal coliform bacteria in streams.   Cattle are one of the sources of fecal coliform bacteria in the Little Arkansas River Basin. High fecal coliform bacteria densities occur at least twice during the year. During the spring when there is considerable rainfall, runoff from cattle-producing areas may result in large amounts of fecal coliform bacteria reaching streams. At the other end of the seasonal cycle, fecal coliform bacteria densities can be high in the dry, late fall and winter months when the cattle congregate near streams.

For both the Halstead and Sedgwick gaging stations, the second year of data collection offered a significant improvement in the sum of squares of error (SSE). The SSE did not improve significantly for either gaging station after 1996 (table 2). Although 2 years of data were sufficient to define this relation at these two gaging stations, this does not mean that 2 years of data collection would be sufficient to define this relation at other sites.

The regression equation for fecal coliform bacteria cannot estimate extreme fecal coliform bacteria densities that sometimes occur as a result of spills or point-source contamination. The estimates from the regression equation are similar to measured densities for smaller bacteria values that may occur as a result of nonpoint-source contamination, such as livestock production. However, if a spill, such as a breached waste lagoon, contributes to the bacteria density, the regression equation cannot estimate this accurately.

Relative percentage differences between measured and estimated loads were 242 percent at the Halstead gaging station and 83.7 percent at the Sedgwick gaging station (Christensen and others, 2000). These large differences are due not only to error in the regression, but also to the amount of error involved in the laboratory analysis of fecal coliform bacteria. Standard methods for analyzing bacteria can have several substantial sources of error such as the length of time before sample analysis, exposure to direct sunlight, temperature during storage, presence of different types of bacteria, and errors in the enumeration of colonies.

Estimated fecal coliform bacteria loads were slightly larger for the Sedgwick gaging station than for the Halstead station (32,400,000 and 30,000,000 billion colonies per year, respectively), and estimated yields were larger for the Halstead station than for the Sedgwick station (68.3 and 43.7 billions of colonies per acre per year, respectively) (fig. 4).

Although samples were collected during a wide range of streamflows, some of the differences between the measured (calculated from manually collected samples) and estimated bacteria loads (fig. 5) could be due to factors other than turbidity. Examples of such factors include manure application, feedlot runoff, sewage-treatment-plant discharges, precipitation characteristics, soil characteristics, and topography.

The dependent variables in equations 1 through 4 were transformed; therefore, consideration must be given to retransformation bias when interpreting the results of the regression analysis. The retransformation has no effect on the form of the equations or on the error associated with the equations. However, retransformation can cause an underestimation of TSS and fecal coliform bacteria loads when adding individual load estimates over a long period of time. Cohn and others (1989), Gilroy and others (1990), and Hirsch and others (1993) provide additional information on the interpreting the results of regression-based load estimates.

DISCUSSION AND CONCLUSIONS

Continuous and periodic monitoring have allowed the identification of trends in turbidity, TSS, and fecal coliform bacteria and the estimation of chemical load transported in the Little Arkansas River. Varying chemical concentrations during high flows have a substantial effect on calculated chemical loads, and concentration data from manual samples often are not available for these conditions. Therefore, continuous monitoring of turbidity for the

Table 2. Sample-size effect on improving sum of squares of error (SSE) for surrogate-based fecal coliform bacteria equation

[R^2, coefficient of determination; SSE, sum of the squares of error;%, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143671)				Little Arkansas River near Sedgwick (station 07144100, fig. 1)			
Calendar year	**Total number of samples**	**R^2**	**SSE**	**Change in SSE (%)**	**Total number of samples**	**R^2**	**SSE**	**Change in SSE (%)**
1995	20	-0.574	75.5	--	18	0.043	94.1	--
1996	42	.578	30.1	-60.1	36	.567	42.6	-54.7
1997	58	.606	28.1	-6.64	50	.593	40.0	-6.10
1998	75	.620	27.1	-3.56	73	.556	43.6	9.00

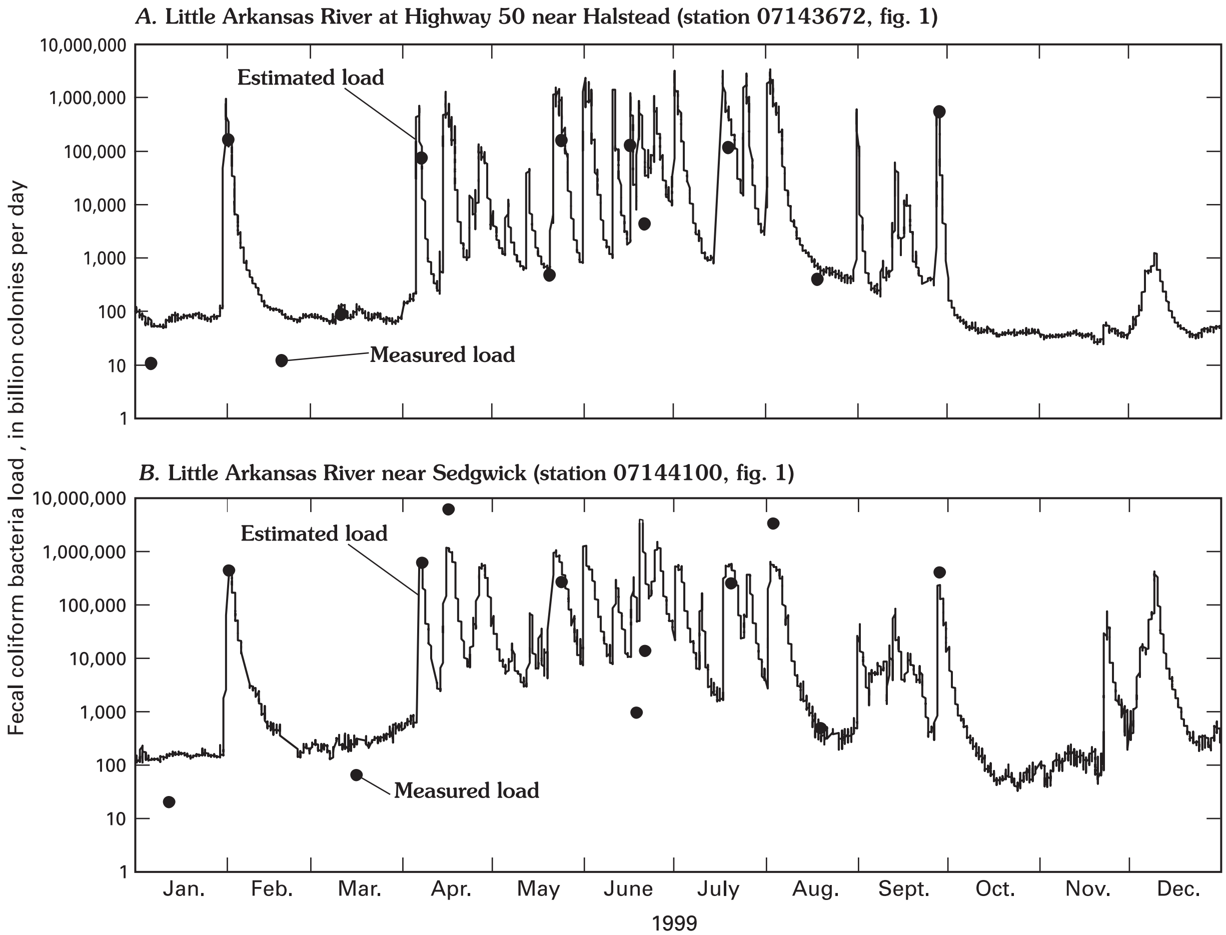

Figure 5. Comparison of measured and estimated fecal coliform bacteria loads for Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

estimation of TSS and fecal coliform bacteria in streamflow may increase the accuracy of load estimates.

The regression-estimated mean daily loads of TSS and fecal coliform bacteria may be more reflective of actual loads than the calculated loads from periodic manually collected data because of the continual nature of the in-stream data. There are few or no gaps in the real-time monitoring of turbidity, except in the case of water-quality monitor malfunction. On the other hand, TSS and bacteria calculated loads from manually collected samples are based on approximately 10 to 15 discrete samples collected throughout the year, and peaks in bacteria densities may have been missed.

In addition to the utility of the regression equations to the city of Wichita for recharge purposes, they also may be useful for calculating total maximum daily loads (TMDL's), which the State of Kansas is mandated to establish for stream segments that have been identified by section 303 (d) of the 1972 Clean Water Act as limited for specific uses because of water-quality concerns. With the development of surrogate relations between continuous turbidity measurements and periodic collection of samples for analysis of TSS and fecal coliform bacteria, a more accurate representation of actual daily loads is probable.

Information on loads and yields may be an indication of which subbasin in which to concentrate efforts with regard to land-resource best-management practices (BMP's). Estimated TSS loads were 460 and 613 million pounds per year for the Halstead and Sedgwick gaging stations, respectively. Estimated fecal coliform bacteria loads were 30,000,000 and 32,000,000 billions of colonies per year. Constituent loads alone are more substantial at the Sedgwick gaging station due to its downstream location and thus higher streamflows. However, when estimated yields are compared, the Halstead subbasin has much higher yields for both constituents. This information can be used by resource managers to prioritize implementation of BMP's.

In addition to the continuous nature of the turbidity data, the data also are available in real time. The timely availability of TSS and fecal coliform bacteria data may be important when considering whole-body contact and recreation criteria for a water body; water suppliers would have timely information to use in adjusting water-treatment strategies; and environmental effects could be assessed in time to prevent adverse effects on fish or other aquatic life.

REFERENCES

Christensen, V.G., Jian, Xiaodong, and Ziegler, A.C., 2000, Regression Analysis and Real-Time Water-Quality Monitoring to Estimate Constituent Loads and Yields in the Little Arkansas River, South-Central Kansas, 1995–99. U.S. Geological Survey Water Resources Investigations Report 00–4126, 36 p.

Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., and Wells, D., 1989, Estimating Constituent Loads. Water Resources Research, v. 25, no. 5, p. 937–942.

Gilroy, E.J., Hirsch, R.M., and Cohn, T.A., 1990, Mean Square Error of Regression-Based Constituent Transport Estimates. Water Resources Research, v. 26, no. 9, p. 2069–2077.

Guy, H.P., 1969, Laboratory Theory and Methods for Sediment Analysis. U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. C1, 58 p.

Hirsch, R.M., Helsel, D.R., Cohn, T.A., and Gilroy, E.J., 1993, Statistical Analysis of Hydrologic Data, *in* Maidment, D.R., ed., Handbook of Hydrology. New York, McGraw-Hill, Inc., p. 17.1–17.55.

Kansas Department of Health and Environment, 1999, List of Water Quality Limited Stream Segments Impacted Predominantly by Nonpoint and Point Sources. accessed March 2, 2000, at URL http://www.kdhe.state.ks.us/befs/303d/table1.htm

Stramel, G.J., 1967, Progress Report on the Ground-Water Hydrology of the Equus Beds Area, Kansas, 1966. Kansas Geological Survey Bulletin 187, part 2, 27 p.

Topping, D.J., Rubin, D.M., Nelson, J.M., Kinzel, P.J., III, and Corson, I.C., 2000, Colorado River Sediment Transport, Part 2. Systematic Bed-Elevation and Grain-Size Effects of Sandy Supply Limitation. Water Resources Research, February 2000, v. 36, no. 2, p. 543.

Watts, K.R., and Stullken, L.E., 1985, Generalized Configuration of the Base of the High Plains Aquifer in Kansas. U.S. Geological Survey Open-File Report 81–344, 1 sheet, scale 1:500,000.

Ziegler, A.C., Christensen, V.G., and Ross, H.C., 1999, Preliminary Effects of Artificial Recharge on Ground-Water Quality, 1995–98—*Equus* Beds Ground-Water Recharge Demonstration Project, South-Central Kansas. U.S. Geological Survey Water-Resources Investigations Report 99–4250, 74 p.

Ziegler, A.C., and Combs, L.J., 1997, Baseline Data-Collection and Quality-Control Protocols and Procedures for the *Equus* Beds Ground-Water Recharge Demonstration Project near Wichita, Kansas, 1995–96. U.S. Geological Survey Open-File Report 97–235, 57 p.